Prediction of speech masking release for fluctuating interferers based on the envelope power signal-to-noise ratio

Jørgensen, Søren; Dau, Torsten

Published in:
Journal of the Acoustical Society of America

Link to article, DOI:
10.1121/1.4708508

Publication date:
2012

Title:

Prediction of speech masking release for fluctuating interferers based on the envelope power signal-to-noise ratio.

Authors: Søren Jørgensen and Torsten Dau.

Abstract:

The speech-based envelope power spectrum model (sEPSM) presented by Jørgensen and Dau [(2011). J. Acoust. Soc. Am. **130**, 1475-1487] estimates the envelope signal-to-noise ratio (SNR_{env}) after modulation-frequency selective processing, which accurately predicts the speech intelligibility for normal-hearing listeners in conditions with additive stationary noise, reverberation, and nonlinear processing with spectral subtraction. The latter condition represents a case in which the standardized speech intelligibility index and speech transmission index fail. However, the sEPSM is limited to conditions with stationary interferers due to the long-term estimation of the envelope power and cannot account for the well known phenomenon of speech masking release. Here, a short-term version of the sEPSM is presented, estimating the envelope SNR in 10-ms time frames. Predictions obtained with the short-term sEPSM are compared to data from Kjems *et al.* [(2009). J. Acoust. Soc. Am. **126** (3), 1415-1426] where speech is mixed with four different interferers, including speech-shaped noise, bottle noise, car noise, and a highly non-stationary cafe noise. The model accounts well for the differences in intelligibility observed for the stationary and non-stationary interferers, demonstrating further that the envelope SNR is crucial for speech comprehension.